Quality of Experience (QoE) and Quality of Service (QoS)

TOPICS*:*

- Introduction
- QoS measures
- Subjective evaluation of voice and video quality
- Objective evaluation of voice, video-telephony, and IPTV quality

1

Quality of Experience (QoE)

- QoE is a term that describes the experience of a user with a particular service.
- The user experience is subjective.
- In networking services the factors that affect QoE can be classified into two groups:
  - – Usability factors
  - – Networking factors

2

Usability factors

- Related to the ease of using a device for accessing a service. Factors include:
  - usability of physical controls
  - ergonomic design and usability
  - attributes of the device
  - the application/service interface itself including its functional format
  - progress feedback (i.e. tones to indicate how the setting up of a service is progressing)
  - mental effort (cognitive load)

Networking factors

Networking factors may be grouped into: responsivity, media fidelity, and availability.

- Responsivity: Network response time, and timely responses to commands.
- Media fidelity: Accuracy of received signal, echo, sound level, speech/video distortion due to errored/lost packets, and jitter.
- Availability: Expresses the percent of time the network is available to transport commands and data.

Quality of service (QoS)

- QoS describes the quality of the underlying network. It is typically expressed in term of three metrics:
  - *end-to-end delay,*
  - *jitter,*
  - *packet loss*.
- These metrics are related to some of the networking factors described above.

QoS metrics for common applications

Tolerance for transport errors	Interactive Delay << 1	Responsive Delay ~ 1s	Timely Delay ~ 10 s	Background Delay >> 10 s
Tolerant	Conversational Voice & video	Voice mail	Streaming Audio & video	Fax
Intolerant	Remote App, Command & Control games	E-commerce Web browsing	IM FTP (foreground)	FTP (background) email

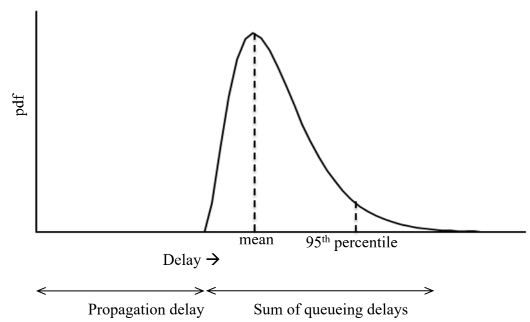
Tolerance for delay

QoS measures: End-to-end delay

- This is the time it takes to deliver a packet from the transmitter to the receiver. It is made up of a *fixed* component and a *variable* component.
- Fixed component: propagation delay, fixed delays induced by transmission systems, and fixed switch processing times.

- Variable component: This is the sum of all variable delays that a packet encounters from the transmitting device to the receiving device. These delays are primarily due to queueing delays in the routers along the packet's path.

QoS measures

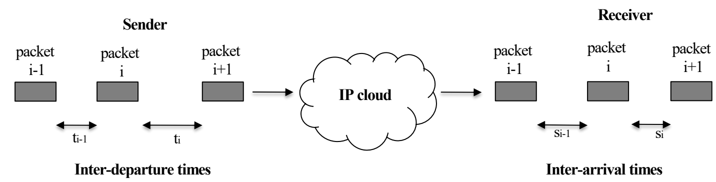


Measures of the end-to-end delay

- The end-to-end delay is measured in terms of its *mean*, but also it can be given as a statistical upper bound in the form of a *percentile*. That is, we say that the end-to-end delay is less than, say 30 ms, 95% of the time.

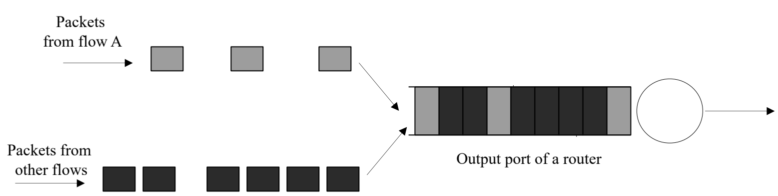
QoS measures: Jitter

- This is the variability of the inter-arrival time of packets arriving at the destination.
- It is important for real-time applications



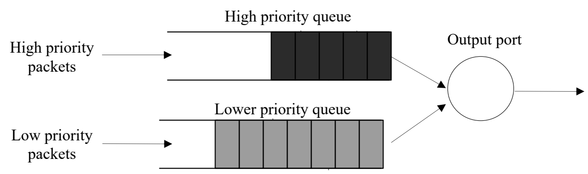
What is the cause of jitter?

- Packets of a flow may be interleaved from other flows in an output buffer, thus elongating their inter-departure time



What is the cause of jitter?

- Packets of a flow may be bunched up when a higher priority queue is served, thus departing back-to-back when their queue is served

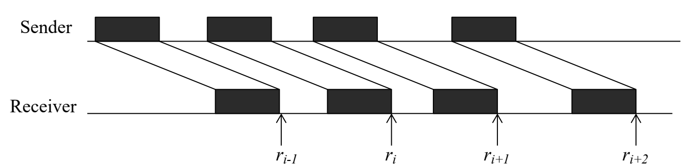


- Jitter is a well understood concept, but there is no agreed upon statistic for measuring it.
- In view of this, various statistics are used. These statistics deal with either the
  - *variability of the inter-arrival times of successive packets at the destination*

  or with the
  - *variability of the one way end-to-end delay*

Variability of the inter-arrival time at the destination

In this case, jitter is expressed as the average of the successive inter-arrival times at the destination. However, more commonly it is expressed as the percentile of the inter-arrival times of the successive packets at the destination. This is an inter-arrival time so that γ% of the inter-arrival times are below it, where γ% could be 95%, 99%, etc.



- Inter-arrivals:

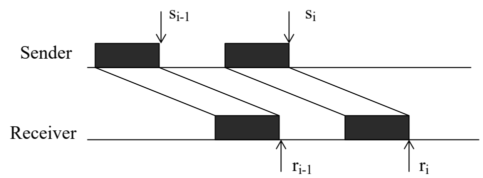
$$j_i = r_i - r_{i-1},\ j_{i+1} = r_{i+1} - r_i,\ j_{i+2} = r_{i+2} - r_{i+1}$$

- Based on this sequence, one can calculate the *average inter-arrival time*, or a *percentile of the inter-arrival times*, such as, the 95th or the 99th percentile.

Variability of the end-to-end delay

- *Inter-packet delay variation (IPDV)*

  It is the average of the difference of the one-way end-to-end delay of successive packets.



$d_{i-1} = r_{i-1}\text{-}s_{i-1},\ d_i = r_i\text{-}s_i$,,

$IPDV_i = d_i\text{-}\ d_{i-1} = (r_i\text{-}s_i)\text{-}(r_{i-1}\text{-}s_{i-1})$

$= (r_i\text{-}r_{i-1})\text{-}(s_i\text{-}s_{i-1})$

- Example: Let d_i=10, 20, 15, 25, 18, 23. Then, $IPDV_i$ = 10, -5, 10, -7, 5. Average: IPDV=13/5.

- *Packet delay variation (PDV)*

  It is defined as the difference between the one way end-to-end delay of each packet, d_i, from the smallest observed one-way end-to-end delay, d_{min}. That is, $d_i\text{-}d_{min}$.
- The PDV is the average of these differences.
- Example: Let d_i=10, 20, 15, 25, 18, 23 (same as above). Then, PDV_i=0, 10, 5, 15, 8, 13, and PDV=51/6.

- *Packet delay variation (PDV)*

  If the smallest end-to-end delay is equal to the propagation delay, then this difference is the sum of all the queueing delays encountered by a packet end-to-end.
- The PDV is then the average of the end-to-end queueing delay, experienced by all packets.

Jitter calculation in RTP

- Let d_i be the one way end-to-end delay, i.e. d_i=(r_i-r_{i-1})–(s_i-s_i-$_1$), where s_i and r_i are the time instances that packet *i* was sent and received, respectively.
- Then the jitter J_i of packet *i* is:

$$J_i = J_{i-1} + (|d_i| - J_{i-1})/16 = (15J_{i-1} + |d_i|)/16$$

Jitter buffer

- The jitter is used to calculate the size of the *jitter buffer*, also known as *de-jitter buffer*.
- A jitter buffer is used at the destination to ensure a continuous play-out of audio and video.
- Packets arrive asynchronously but their contents has to be played synchronously.

Fixed vs adaptive jitter buffers

- Jitter buffers can be fixed or adaptive.
- Fixed jitter buffers do not change their buffer size
- Adaptive ones change their buffer size as a function of how fast or slowly packets arrive.

- The size of a jitter buffer is expressed either in terms of its storage size or in terms of the total time it takes to drain a full buffer (max storage time), through the expression:

*read rate * max. storage time = size of buffer*

- Ex: G.711 codec, read rate 64 Kbps, and assuming a maximum storage time of 30 msec, the size of the buffer is 240 bytes.
- Typically, the size of a jitter buffer is expressed in terms of *max. storage time*.

Jitter buffer for streaming applications

- Sizing the jitter buffer is simple for streaming applications (video-on-demand, IPTV, radio).
- Before the play-out process starts playing, it buffers 3 to 4 seconds worth of voice data in order to absorb delays in the packet arrivals.
- Persistent delays causes the play-out process to run out of data. In this case, it stops, re-buffers, and restarts.

Jitter for VoIP

- The delay in the jitter buffer should not be large so that the end-to-end delay is significantly affected.
- On the other hand, if the jitter buffer is small, it may run out of data due to the packet inter-arrival variability.
- The speech gaps that occur when the buffer runs out of data maybe hidden by different concealment techniques that replace the data that are lost or not received in time.

- If a VoIP packet experiences the minimum delay, d_{min}, then its voice data could suffer the maximum delay in the jitter buffer b_{max}.
- The sum $d_{min} + b_{max}$ should be equal to the maximum delay d_{max} plus the minimum buffer delay b_{min}. That is:

$$d_{min} + b_{max} = d_{max} + b_{min}.$$

Or,

$$b_{max} - b_{min} = d_{max} - d_{min}.$$

- That is, the range of the buffer delay is equal to the range of the one way end-to-end delay,

- Example: Let 10, 20, 15, 25, 18, 23, be the one-way end-to-end delay of 6 packets. Then, the range of the PDV is 25-10=15.
- Assuming that b_{min}=0, then the maximum delay applied in the jitter buffer is b_{max} equal to 15.
- Typically the range of the one-way end-to-end delay may be replaced by the 99.9^{th} percentile of the PDV.
- This delay is typically referred to as the *jitter delay*.

- Cisco recommends that the actual size of the jitter buffer should be 1.5 to 2 times the jitter delay.
- Since traffic conditions change, the jitter is typically calculated over a short window of around 100 msec, and the size of the jitter buffer is fixed accordingly.
- Therefore, the jitter buffer is continuously adjusted, but it stays fixed for each window of time.

Operation of the VoIP jitter buffer

- The first packet that arrives finds the jitter buffer empty and it is delayed by the jitter delay, say 30 msec. Then, the voice data will start getting delivered to the codec.
- Packets that arrive when the jitter buffer is not empty are queued. If a packet finds the jitter buffer empty, it is delayed by 30 msec as above.
- Out of order packets are re-ordered and consumed if they have arrived on time. Otherwise, they are lost packets.
- If the buffer overflows, the packets will get lost resulting in a choppy voice signal.

QoS measures: packet loss

- Packet loss is expressed as the percent of packets that are lost.
  - Packets are typically lost in the buffers of a router when congestion arises.
  - Also when they are transmitted over a link if there is noise on the link. (Very low prob.)
  - Finally, a packet is lost at the IP and TCP level if errors are detected.

Not all packets are equal ..

- Depending on the application, when a packet is lost the user may or may not perceive this at the application level.
- Not all lost packets may have the same impact on quality. Example, a packet lost belonging to an I-frame as opposed to a B or P frame.

QoS requirements for some services

QoS requirements	jitter	End-to-end delay	Packet loss
VoiP	30 msec	≤150 msec	≤ 1%
Interactive video	30 msec	≤150 msec	≤ 1%
Streaming video	No significant requirement	up to 4-5 sec	≤ 5%
Peer-to-peer	No significant requirement	No significant requirement	No significant requirement

Jitter is defined here as the average difference of the end-to-end delay of successive packets. Jitter is part of the e2e delay budget.

Subjective evaluation of voice and video quality

- The QoE of a user is subjective and it is typically measured using subjective and objective assessment methods.
  - Subjective: A number of listeners are asked to rate their perceived QoE.
  - Objective: Software tools are used to estimate QoE, based on measured QoS and other parameters.

Subjective evaluation of voice quality

- Mean opinion score (MOS) is used to rate the perceived quality of received audio after compression and/or transmission.
- A number of listeners rate the heard quality of experience of test sentences read aloud by both male and female speakers over the communications medium being tested, and the results are averaged out.

- Typical sentences used are:
  - You will have to be very quiet.
  - There was nothing to be seen.
  - They worshipped wooden idols.
  - I want a minute with the inspector
  - Did he need any money?
- A listener gives each sentence a rating from 1 to 5
  - 5 (Excellent) - like face-to-face conversation.
  - 4 (Good) – perceived imperfections, but sound is clear.
  - 3 (Fair) – slightly annoying.
  - 2 (Poor) – annoying.
  - 1 (Bad) – very annoying.

MOS values

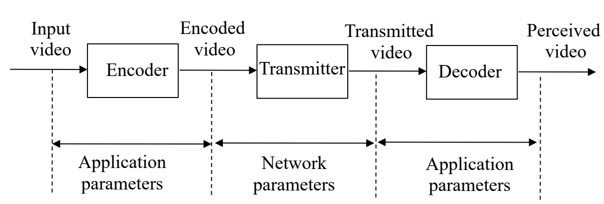
- A MOS value of 4.0 to 4.5 is referred to as toll-quality and causes complete satisfaction.
- This is the normal value of the telephone system and many VoIP services aim at it.
- Values dropping below 3.5 are termed unacceptable by many users.

MOS values for some codecs

- Pulse Code Modulation (PCM): 64 Kbps data stream and has a MOS value of 4.3.
- *Internet Low Bit Rate Codec (iLB):* 15.2 Kbps data stream with a MOS value of 4.14.
- *G.729:* 8 Kbps data stream, MOS=3.92.
- *G.729A:* 8 Kbps data stream, MOS=3.7.
- *G.723.1:* 6.4 Kbps (MOS 3.9) and 5.3 Kbps (MOS 3.62).

Subjective evaluation of video quality

- Parameters that affect the quality of perceived video



MOS calculation

- A number of video sequences and the setting of the application and networking parameters is also selected.
- A number of viewers is used and each viewer rates numerically the perceived quality of the video.
- The average value is then calculated.

Testing methods

- Various testing methods have been developed, such as:
  - *Double Stimulus Impairment Scale* (DSIC) method
  - *Double-Stimulus Continuous Quality Scale* (DSCQS)
  - *Single-Stimulus Continuous Quality Evaluation* (SSCQE),

Double stimulus impairment scale (DSIC) method

- The viewer is presented with an unimpaired reference video and then with the same video impaired. Then the viewer is asked to vote on the second video having in mind the first video.
- A test session lasts up to half an hour

The double-stimulus continuous quality-scale (DSCQS) method

- Viewers are asked to view a pair of video sequences, the reference video sequence and the impaired video sequence, without being told which is which. Each video sequence is about 10 sec, and it is shown twice.
- Then viewers assess the quality of each video sequence in the pair. The difference between the two scores is then used to quantify changes in the quality.

Objective evaluation of audio and video quality

- Subjective voice and video quality testing is a reliable method for evaluating subjective quality. However, it is time-consuming, expensive, and requires special assessment facilities.
- Also, only a small sample of test conditions may be presented for assessment in a single test.
- In view of this, objective means of assessing quality have been developed

- The quality of voice and video can be measured a) prior to transmission of the signal, b) inside the network during the transmission, and, c) at the receiving device.
- The measurements obtained can be used for in-service quality monitoring and management, codec optimization, codec selection, and planning of networks and terminals for services such as VoIP, IPTV, video streaming, and mobile TV.

Evaluation of voice quality – The E-model

- The output from the E-model is a scalar quality rating value, referred to as the *R-value*. R-values below 50 are not recommended.

R-values	Speech transmission quality	User satisfaction
90≤*R*<100	Best	Very satisfied
80≤*R*<90	High	Satisfied
70≤*R*<80	Medium	Some users dissatisfied
60≤*R*<70	Low	Many users dissatisfied
50≤*R*<60	Poor	Nearly all users dissatisfied

Some R-values

Service/network scenario	*R*-value
ISDN subscriber to ISDN subscriber, local connection	94
Analogue PSTN subscriber to analogue PSTN subscriber, 20 ms delay (average echo path losses; no active echo control)	82
Mobile subscriber to analogue PSTN subscriber as perceived at mobile side	72
Mobile subscriber to analogue PSTN subscriber as perceived at PSTN side	64
Voice over IP connection using G.729A with voice activity detection and 2% packet loss	55

Mapping to MOS values

- The *R*-values can be mapped into MOS using the following equation:

$$MOS(R) = \begin{cases} 1, & R < 0 \\ 1 + 0.035R + R(R-60)(100-R)7x10^{-6}, & 0 < R < 100 \\ 4.5, & R > 100 \end{cases}$$

Calculation of R-value

R=Ro −Is −Id −Ie-eff +A,

- *Ro* is the basic signal-to-noise ratio,
- *Is* a combination of the impairments which occur more or less simultaneously with the voice signal.
- *Id* represents the impairments caused by delay.
- *Ie-eff* represents impairments introduced by low bit-rate codecs and packet loss.
- *A* is the advantage factor that allows for the compensation of impairments because of access-related advantages.

A simplified expression

$$R=93.2 - Id - Ie\text{-}eff + A$$

- *A* is typically a constant
- If the delays due to echo cancellation are zero, *Id* is approximated by the end-to-end delay
- *Ie-eff* is a function of packet loss and some other parameters associated with each codec.

Evaluation of video-telephony quality

- The model consists of the following three functions:
  - a) a speech quality estimation using a simplified version of the E-model;
  - b) a video quality estimation which depends on application parameters, and,
  - c) a multimedia quality integration function.

Evaluation of IPTV quality

- A QoE expression for IPTV has been proposed that consists of a network QoE, (QoE_N), and a user QoE (QoE_U).
  - QoE_N is inversely proportional to the end-to-end delay, jitter and packet loss.
  - QoE_U is a function of the synchronization time between audio and video (sync), the zapping time (zap), and *Video Quality* (VQ).